

SCIENTIFIC REPORTS

OPEN

A hybrid stochastic model of folate-mediated one-carbon metabolism: Effect of the common C677T *MTHFR* variant on *de novo* thymidylate biosynthesis

Karla Misselbeck^{1,2}, Luca Marchetti¹, Martha S. Field³, Marco Scotti⁴, Corrado Priami^{1,2} & Patrick J. Stover³

Folate-mediated one-carbon metabolism (FOCM) is an interconnected network of metabolic pathways, including those required for the *de novo* synthesis of dTMP and purine nucleotides and for remethylation of homocysteine to methionine. Mouse models of folate-responsive neural tube defects (NTDs) indicate that impaired *de novo* thymidylate (dTMP) synthesis through changes in SHMT expression is causative in folate-responsive NTDs. We have created a hybrid computational model comprised of ordinary differential equations and stochastic simulation. We investigated whether the *de novo* dTMP synthesis pathway was sensitive to perturbations in FOCM that are known to be associated with human NTDs. This computational model shows that *de novo* dTMP synthesis is highly sensitive to the common *MTHFR* C677T polymorphism and that the effect of the polymorphism on FOCM is greater in folate deficiency. Computational simulations indicate that the *MTHFR* C677T polymorphism and folate deficiency interact to increase the stochastic behavior of the FOCM network, with the greatest instability observed for reactions catalyzed by serine hydroxymethyltransferase (SHMT). Furthermore, we show that *de novo* dTMP synthesis does not occur in the cytosol at rates sufficient for DNA replication, supporting empirical data indicating that impaired nuclear *de novo* dTMP synthesis results in uracil misincorporation into DNA.

Perturbations in folate-mediated one-carbon metabolism (FOCM) are associated with numerous pathologies including neural tube defects (NTDs)¹, stroke², colorectal and other types of cancer^{3–5}. Furthermore, enzymes in FOCM have been successful targets for the development of antineoplastic pharmaceutical agents including 0020methotrexate and 5-fluorouracil⁶. FOCM in the cytoplasm is composed of three interconnected biosynthetic pathways, which include *de novo* thymidylate (dTMP) synthesis, *de novo* purine synthesis and homocysteine remethylation to methionine (Fig. 1). The FOCM network is sensitive to nutritional status for several vitamins that serve as enzyme cofactors (folate, riboflavin, vitamin B6 and vitamin B12) and genetic factors (coding and expression variants in folate-dependent enzymes) that can alter network outputs, including DNA synthesis, DNA repair and chromatin methylation^{7–9}. Understanding the molecular basis of disease etiology has been limited by the ability to ascribe specific FOCM pathways and their biomarkers to clinical outcomes, because the pathways of FOCM are tightly interconnected⁸. FOCM complexity is manifest by: (a) competition among the pathways for a limiting pool of folate cofactors¹⁰, (b) long-range and indirect regulatory processes, (c) formation of multi-enzyme complexes, (d) cellular compartmentation, (e) interactions with other metabolic pathways, (f) nutritional status (g) penetrant genetic variants^{9,11}. Mathematical models have been developed to assess this

¹The Microsoft Research - University of Trento Centre for Computational and Systems Biology (COSBI), Piazza Manifattura, 1, 38068, Rovereto (TN), Italy. ²Department of Mathematics, University of Trento, Trento, Italy. ³Division of Nutritional Sciences, Cornell University, Ithaca, New York, 14853, USA. ⁴GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsterbrook Weg 20, 24105, Kiel, Germany. Karla Misselbeck and Luca Marchetti contributed equally to this work. Corrado Priami and Patrick J. Stover jointly supervised this work. Correspondence and requests for materials should be addressed to C.P. (email: priami@cosbi.eu) or P.J.S. (email: pjs13@cornell.edu)

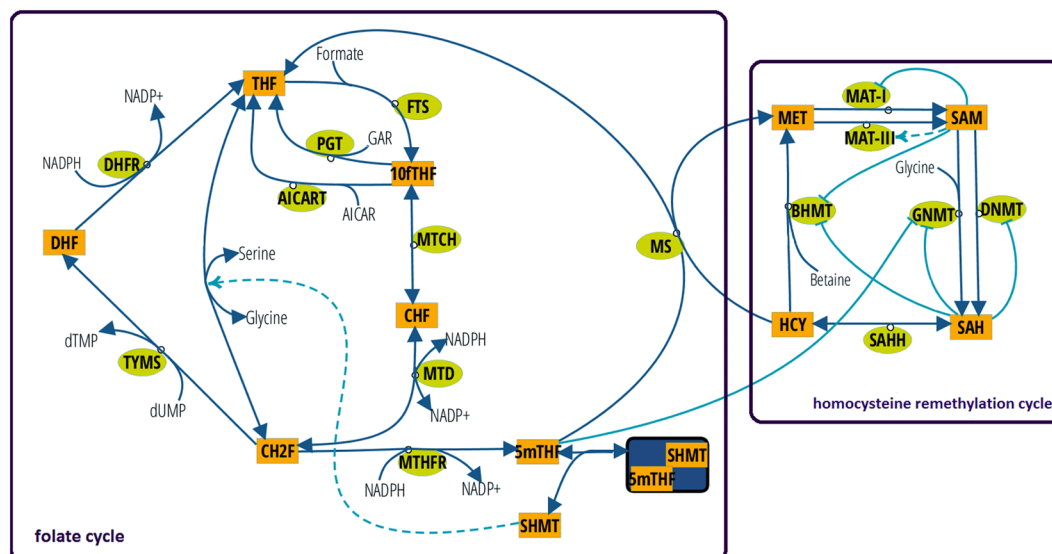


Figure 1. The reaction-based specification of the model according to the notation introduced in Gostner *et al.*⁶¹. Rectangles identify model variables, non-boxed substrates are model constants, green circles identify enzymes, dark blue arcs identify matter transformation, and light blue arcs identify regulatory events (dotted lines indicate activations and solid lines indicate inhibitions). The purple boxes indicate reactions and variables associated with the folate cycle and the homocysteine remethylation cycle, respectively.

complexity and gain an understanding of the cause-and-effect relationships that regulate FOCM functioning in health and disease. The overall goal is to provide an understanding of function of the entire system *in silico* that can be used to accelerate discovery and guide the design of biological experimentation.

Early models of FOCM quantified steady-state concentrations of folates, and predicted how methotrexate and 5-fluorouracil affect the rates of *de novo* purine and thymidylate biosynthesis in L1210 cells^{12–15}. These models considered FOCM reactions as occurring in a common cellular compartment, described the reactions in terms of Michaelis-Menten kinetics, and investigated biological consequences of inhibiting *de novo* purine and dTMP biosynthesis. Other models have been constructed to simulate the homocysteine remethylation in liver^{16–18}. Using data from experimental animals, non-trivial mathematical functions were adopted to address long-range inhibition and activation processes involving folates and other metabolites that cannot be translated into simple Michaelis-Menten equations¹⁷. These models can reproduce changes in metabolite concentrations in response to nutritional deficiencies and the effects of gene variants, thus corroborating human clinical and epidemiological data. These models have also integrated other pathways such as glutathione metabolism and polyamine synthesis, studying the influence of indirect effects beyond FOCM^{19,20}.

These existing FOCM models describe the continuous flux of metabolite concentrations in terms of ordinary differential equations (ODEs) with time as an independent variable^{13,20}. Sensitivity analysis is used to summarize the effects on metabolite concentrations and reaction velocities in response to changes in inputs or enzyme activities²¹. Here we present a hybrid stochastic model for simulating the FOCM dynamics where state-of-the-art deterministic simulation (based on ODEs) has been coupled with exact stochastic simulation to assess metabolite variabilities in the FOCM network at steady state. A detailed description of this combined approach can be found in Methods. The deterministic approach used in isolation can only provide a rough estimate of FOCM model dynamics, because the deterministic approach is limited when enzyme substrates, such as folate cofactors, are present in the cells at low micromolar concentrations, and because reactions within the network occur randomly at discrete time points. FOCM is expected to exhibit variability (i.e. stochasticity) in its behavior²². Capturing system stochasticity is essential when substrate concentrations are low and limiting, but requires consideration of molecules as discrete entities, rather than describing concentrations as continuous variables through ODEs²³. Simulation strategies that combine both deterministic and stochastic approaches can give a more accurate and more detailed understanding of FOCM network functioning and stability. In contrast to approaches based solely on deterministic simulation, these studies can be used to assess the contributions of factors such as genetic variation and nutritional status on the stochastic behavior of individual pathways within the network, thereby aiding in establishing which system inputs (i.e. nutrition) and outputs (i.e. biomarkers) are most closely associated with human health outcomes.

Dysregulation of the partitioning of one-carbon units in the form of 5,10-methylenetetrahydrofolate (CH₂F) cofactors between the *de novo* dTMP biosynthesis and homocysteine remethylation pathways is believed to underlie FOCM-associated pathologies including NTDs (Fig. 1). A common variant of *MTHFR*, the C677T polymorphism, has been associated with numerous pathologies including birth defects, cancer, cardiovascular events, and other pathologies²⁴. *MTHFR* catalyzes the FADH-dependent, irreversible conversion of CH₂F to 5mTHF, which commits folate cofactors away from dTMP synthesis and towards homocysteine remethylation in the cytosol (Fig. 1). The variant results from an alanine to valine substitution in the protein that decreases *MTHFR*

Glycine	THF	10fTHF	CHF	CH2F	DHF	5mTHF			HCY	MET	SAM	SAH
	%	%	%	%	%	%free	%bound	%total	μM	μM	μM	μM
0 mM	0.19	36.17	7.25	2.17	0.04	10.31	48.87	54.18	2.72	30.58	99.99	17.23
1 mM	0.45	38.25	7.55	2.02	0.04	8.78	42.90	51.68	3.44	42.09	54.44	50.55
2 mM	0.74	39.37	7.70	1.91	0.04	7.97	42.26	50.23	3.48	42.48	50.96	53.60
5 mM	1.86	40.80	7.87	1.72	0.03	6.71	41.01	47.73	3.52	42.83	46.49	57.67
10 mM	5.31	40.40	7.71	1.52	0.03	5.56	39.48	45.04	3.57	43.04	42.40	61.51

Figure 2. Steady states for five different values of glycine (folate cycle in % and the homocysteine remethylation cycle in μM). Where possible, a trend arrow is provided on the right to show the experimental outcome observed in Herbig *et al.*²⁸. All the trends are consistent with literature (green arrows) except for the total % of 5mTHF in the glycine range 5–10 mM (red arrow).

activity by decreasing its affinity for the FADH cofactor. Such substitution affects enzyme stability and hence the partitioning of folates between dTMP synthesis and homocysteine remethylation^{25, 26}. Decreased MTHFR activity resulting from the polymorphism decreases 5mTHF synthesis, leading to impaired homocysteine remethylation and elevated serum homocysteine²⁴. The 677 T variant is also associated with a redistribution of cellular folate cofactors; 5mTHF is the predominate form of folate in red blood cells in *MTHFR* 677CC carriers, whereas 10fTHF is the predominate form of folate in *MTHFR* 677TT carriers^{27–29}. 10fTHF is less chemically stable than 5mTHF, and the *MTHFR* 677TT variant is associated with lower folate status³⁰ and higher folate requirement³¹. Recent studies suggest that the contribution of the *MTHFR* variant to NTD risk is due to its impact on cellular folate status, rather than impaired homocysteine remethylation³². Likewise, mouse models of NTDs indicate that impaired dTMP synthesis, and not homocysteine remethylation, cause folate-responsive NTDs^{33–35}. Reed *et al.* investigated the consequences of the *MTHFR* C677T polymorphism, assuming 70% enzyme activity for heterozygote and 30% enzyme activity for homozygote, in comparison to CC homozygotes (which was set to 100% activity) using parameters for folate monoglutamates, which are not the physiological form of folate cofactors in cells. Under these conditions, the variant allele decreased concentrations of 5mTHF and SAM and increased the concentrations of homocysteine, SAH, and rates of dTMP and purine biosynthesis²¹, which is inconsistent with current understanding of biochemical changes associated with NTD risk. The effect on the redistribution of folate cofactors towards 10fTHF that is associated with the 677 T variant, or its impact on other pathways within the network, was not reported²¹.

Here, we studied the partitioning of CH2F, a cofactor for both homocysteine remethylation and *de novo* dTMP biosynthesis^{36–38}, and the effects of known genetic and nutritional variables that impact movement of CH2F through the network. Existing models are limited by adopting kinetic parameters determined from the use of folate monoglutamate substrates^{19, 20}. Folate polyglutamates are the functional form of folate cofactors in cells and have much higher affinity for their respective FOCM enzymes than the corresponding monoglutamate forms of folate^{39, 40}. Therefore, we updated the parameters in the deterministic model to include the physiologically-relevant polyglutamate forms of folate cofactors and demonstrate that it faithfully recapitulates existing data in the literature (see Methods, Model Validation and Fig. 2). The decisions for selecting individual enzyme kinetic parameters for this model were driven by: 1) available data for physiologically relevant polyglutamate forms of the folate cofactors derived from characterization of mammalian enzymes, and 2) data from human models, specifically L1210 cells, because of the richness and quality of the data used to derive kinetic parameters. Given the high conservation of folate enzymes among mammals, our model could be applied to mammalian systems in general, even though we are not proposing a completely homogeneous model with respect to species.

We were able to identify key nodes in the network of Fig. 1 that exhibit high degrees of stochastic behavior, including the influence of nutrient status and genetic variation on stochasticity through simulations. We explored the impact of the *MTHFR* C677T polymorphism and its interaction with folate status on partitioning of CH2F within the network, including its impact on *de novo* dTMP biosynthesis to understand the etiology of NTDs. The results of the computational model provide evidence that the rates of *de novo* dTMP synthesis as currently modeled in the cytosol are insufficient to support DNA synthesis in S-phase in mammals, accounting for uracil misincorporation into DNA that occurs in folate deficiency and in mouse models of NTDs.

Results

The effect of folate status and the *MTHFR* polymorphism on pathways affecting NTD risk. In the current model, MTHFR activity was decreased to model the effect of the *MTHFR* C677T polymorphism. In addition, a two-fold increase in MTHFR activity was modeled to examine whether there was a dose-response relationship between MTHFR activity and various readouts of the network. This model shows that 5mTHF levels decrease as MTHFR activity decreases, reflecting the effects of the *MTHFR* C677T polymorphism (Table 1). The current model also recapitulates the biological observation that decreased MTHFR activity results in accumulation of 10fTHF and THF (Table 1). Inclusion of 10fTHF in the folate distribution is a strength of the current model in that it allows for estimation of the effect that perturbations to the system have on accumulation of this unstable form of folate, which likely accounts for the decreased folate status linked to NTDs in carriers of the polymorphism^{30, 32}.

%	THF	10fTHF	CHF	CH2F	DHF	5mTHF		
						free	bound	total
MTHFR x2	0.25	14.43	2.86	0.79	0.01	33.45	48.20	81.65
MTHFR x1 – CC	0.70	39.24	7.69	1.93	0.04	8.07	42.34	50.42
MTHFR x0.7 – CT	1.59	48.76	9.45	2.18	0.04	3.37	34.61	37.98
MTHFR x0.5	3.13	54.53	10.51	2.32	0.04	1.86	27.60	29.46
MTHFR x0.3 – TT	8.12	58.72	11.26	2.40	0.04	0.90	18.56	19.46

Table 1. Distribution of folate (in percentage of total folate) for different levels of MTHFR activity (ranging from 2x to 0.3x of wild type). CC, CT and TT refer to the C667T polymorphism.

μM	THF	10fTHF	CHF	CH2F	DHF	5 mTHF			Unbound SHMT	HCY	MET	SAM	SAH
						free	bound	total					
MTHFR x2	0.04	2.58	0.51	0.14	0.00	5.97	8.60	14.57	0.40	3.10	37.99	78.41	31.01
MTHFR x1 – CC	0.12	7.00	1.37	0.34	0.01	1.44	7.56	9.00	1.44	3.47	42.44	51.34	53.26
MTHFR x0.7 – CT	0.28	8.70	1.69	0.39	0.01	0.60	6.18	6.78	2.82	3.67	43.00	34.87	68.98
MTHFR x0.5	0.56	9.73	1.88	0.41	0.01	0.33	4.93	5.26	4.07	3.78	42.59	26.72	77.42
MTHFR x0.3 – TT	1.45	10.48	2.01	0.43	0.01	0.16	3.31	3.47	5.69	3.89	41.87	20.47	84.29

Table 2. Concentrations of model variables (in μM) for different levels of MTHFR activity (ranging from 2x to 0.3x of wild type). CC, CT and TT refer to the C667T polymorphism.

μM/h	FTS	MTCH		MTD		MTHFR	MTR	Unbinding of SHMT and 5mTHF
		10fTHF → CHF	CHF → 10fTHF	CHF → CH2F	CH2F → CHF			
MTHFR x2	13,101.4	332,703.8	330,313.4	37,733.2	35,342.8	26.8	26.8	17,035.2
MTHFR x1 – CC	23,582.5	756,270.3	744,706.4	89,938.7	78,374.9	21.6	21.6	14,964.5
MTHFR x0.7 – CT	31,430.7	884,173.6	864,936.0	106,222.3	86,984.7	16.0	16.0	12,231.1
MTHFR x0.5	36,054.6	954,589.4	930,806.0	115,378.8	91,595.3	11.7	11.7	9,754.6
MTHFR x0.3 – TT	39,755.7	1,002,680.1	975,243.2	121,653.1	94,216.2	7.1	7.1	6,559.7

Table 3. Fluxes of the reactions catalyzed by the enzymes FTS, MTCH, MTD, MTHFR and MTR and the binding of 5mTHF and SHMT for different levels of MTHFR activity (ranging from 2x to 0.3x of wild type). CC, CT and TT refer to the C667T polymorphism.

SAM and SAH levels vary markedly with changes in MTHFR activity (SAM levels decrease and SAH levels increase by more than 50% when comparing the “CC” model to the “TT” model), with a SAM/SAH ratio around one being achieved in CC homozygotes compared to 0.24 in TT homozygotes (Table 2). Although methylation potential, otherwise known as the SAM/SAH ratio, changes markedly with varying MTHFR activity, homocysteine concentrations appear relatively insensitive to changes in MTHFR in this model (Table 2), inconsistent with known effects of the *MTHFR* TT genotype in elevating serum homocysteine levels⁴¹. However, in this model the lack of elevation in homocysteine due to the *MTHFR* C667T polymorphism reflects that the model represents a closed system leading to intracellular conversion of cellular homocysteine to SAH as opposed to export of homocysteine into the circulation (Table 2).

The activity of each enzyme in the FOCM network as predicted by the current model indicates that accumulation of 10fTHF resulting from decreased MTHFR activity is due to two factors: (1) an increased flux through both the 10fTHF synthetase activity leading to increased synthesis (FTS), and (2) an increased flux through the cyclohydrolase/dehydrogenase activity of MTHFD1 which converts CH2F to CHF to 10fTHF (MTCH, MTD activities, respectively, Table 3). Interestingly, the accumulation of 10fTHF does not affect flux through the enzymes that use 10fTHF as a co-factor for *de novo* purine synthesis (PGT, AICART; Table 4). This is consistent with empirical experimental findings that increasing cellular levels of 10-formyltetrahydrofolate dehydrogenase, which consumes 10fTHF, do not affect *de novo* purine synthesis⁴². Flux through the *de novo* dTMP synthesis pathway increases with decreasing MTHFR activity, consistent with empirical studies indicating these two pathways compete for CH2F (Table 4, columns DHFR, TYMS and SHMT)⁴³. These findings are in agreement with empirical data showing that the TT polymorphism results in an increase in CH2F available for dTMP synthesis as indicated by isotope tracer studies in humans²⁶.

$\mu\text{M/h}$	PGT	AICARFT	DHFR	TYMS	SHMT		Binding of SHMT and 5mTHF
					CH ₂ F \rightarrow THF	THF \rightarrow CH ₂ F	
MTHFR x2	4,406.3	6,304.6	113.5	113.5	2819.8	569.7	17,035.2
MTHFR x1 - CC	5,268.8	6,749.8	263.4	263.4	15,646.6	4,367.7	14,964.5
MTHFR x0.7 - CT	5,388.7	6,804.3	295.0	295.0	31,989.8	13,063.2	12,231.1
MTHFR x0.5	5,442.7	6,828.4	312.3	312.3	47,127.3	23,667.8	9,754.6
MTHFR x0.3 - TT	5,475.7	6,843.0	322.1	322.1	66,523.2	39,415.5	6,559.7

Table 4. Fluxes of the reactions catalyzed by the enzymes PGT, AICARFT, DHFR, TYMS, SHMT and the unbinding of 5mTHF and SHMT for different levels of MTHFR activity (ranging from 2x to 0.3x of wild type). CC, CT and TT refer to the C667T polymorphism.

$\mu\text{M/h}$	BHMT	MAT-I	MAT-III	GNMT	DNMT	SAHH	
						SAH \rightarrow HCY	HCY \rightarrow SAH
MTHFR x2	145.9	110.3	62.4	119.6	127.4	264.6	91.9
MTHFR x1 - CC	161.0	123.0	59.6	370.9	87.2	285.2	102.6
MTHFR x0.7 - CT	168.3	127.9	56.3	593.3	59.6	292.4	108.2
MTHFR x0.5	172.1	129.3	54.4	723.1	45.6	295.2	111.5
MTHFR x0.3 - TT	175.5	129.8	52.7	830.4	34.7	297.1	114.5

Table 5. Fluxes of the reactions catalyzed by the enzymes MTR, BHMT, MAT-I, MAT-III, GNMT, DNMT and SAHH for different levels of MTHFR activity (ranging from 2x to 0.3x of wild type). CC, CT and TT refer to the C667T polymorphism.

%		THF	10fTHF	CHF	CH ₂ F	DHF	5mTHF		
							free	bound	total
low folate	CC	1.27	38.48	7.46	1.65	0.03	4.98	46.13	51.11
	TT	4.24	49.75	9.45	1.68	0.03	0.93	33.91	34.84
replete folate	CC	0.70	39.24	7.69	1.93	0.04	8.07	42.34	50.42
	TT	8.12	58.72	11.26	2.40	0.04	0.90	18.56	19.46

Table 6. Distribution of folate (in percentage of total folate) for replete (18 μM) and low (9 μM) levels of total folate and for the CC and TT case of the C667T MTHFR polymorphism.

5mTHF binds to and is a potent inhibitor of SHMT and glycine *N*-methyltransferase (GNMT). Decreased MTHFR activity, which lowers cellular 5mTHF levels (Table 2), increases the flux through the SHMT1-catalyzed reaction in the direction of serine catabolism to glycine (Table 4). This reflects the depletion of intracellular 5mTHF as well as glycine, resulting from increased GNMT activity that results from 5mTHF depletion (Table 5). Methionine synthase (MTR) flux is highly sensitive to MTHFR genotype reflecting its dependence on availability of its substrate 5mTHF that is generated by MTHFR (Table 3).

The MTHFR C677T variant affects both CH₂F partitioning (between homocysteine remethylation and dTMP biosynthesis) and intracellular folate concentrations; the 677 T variant lowers intracellular folate levels. Therefore, the impact of this variant on FOCM was modeled at two different levels of folate (Tables 6–10) (folate replete conditions, 18 μM , and low-folate conditions, 9 μM) to understand how the variants function within the FOCM network as a function of folate cofactor availability. The results demonstrate that the changes in the percentage of 5mTHF (as well as other major one-carbon forms of folate) are more pronounced in the TT genotype than in the CC genotype when cellular folate levels are decreased (Table 6). The percentage of 5mTHF in CC homozygous does not change in the folate replete and deficiency states, whereas the accumulation of 5mTHF in TT homozygotes differs between the deficient and replete states (Table 6).

Fluxes through FOCM pathways are affected by both the MTHFR C677T polymorphism and folate levels. The most sensitive pathways to folate deficiency are the MTCH and MTD activities of MTHFD1, MTHFR, MTR, DNMT, DHFR, and TYMS (Tables 8–10, row showing absolute flux differences between folate levels). Flux through the dTMP synthesis pathway (DHFR and TYMS) is highly sensitive to folate status for both the MTHFR CC and TT genotypes (Table 9), with the TT homozygotes being the most sensitive. Flux through GNMT was

μM		THF	10fTHF	CHF	CH2F	DHF	5 mTHF			Free SHMT	HCY	MET	SAM	SAH
							free	bound	total					
low folate	CC	0.11	3.43	0.67	0.15	0.00	0.44	4.12	4.56	2.55	3.73	42.84	30.34	73.61
	TT	0.38	4.44	0.84	0.15	0.00	0.08	3.03	3.11	10.00	3.95	41.35	17.43	87.79
replete folate	CC	0.12	7.00	1.37	0.34	0.01	1.44	7.56	9.00	1.44	3.47	42.44	51.34	53.26
	TT	1.45	10.48	2.01	0.43	0.01	0.16	3.31	3.47	5.69	3.89	41.87	20.47	84.29

Table 7. Concentrations of model variables (in μM) for replete (18 μM) and low (9 μM) levels of total folate and for the CC and TT case of the C667T MTHFR polymorphism.

$\mu\text{M/h}$		FTS	MTCH		MTD		MTHFR	MTR
			10fTHF \rightarrow CHF	CHF \rightarrow 10fTHF	CHF \rightarrow CH2F	CH2F \rightarrow CHF		
CC	replete folate	23,582.5	756,270.3	744,706.4	89,938.7	78,374.9	21.6	21.6
	low folate	22,583.6	427,359.0	415,960.1	48,054.6	36,655.7	13.8	13.8
absolute difference (% of replete folate)		4.2	43.5	44.1	46.6	53.2	36.5	36.5
TT	replete folate	39,755.7	1,002,680.1	975,243.2	121,653.1	94,216.2	7.1	7.1
	low folate	33,618.2	529,771.1	507,690.2	59,371.2	37,290.2	4.2	4.2
absolute difference (% of replete folate)		15.4	47.2	47.9	51.2	60.4	41.2	41.2

Table 8. Fluxes of the reactions catalyzed by the enzymes FTS, MTCH, MTD, MTHFR and MTR for replete (18 μM) and low (9 μM) levels of total folate and for the CC and TT case of the C677T MTHFR polymorphism.

$\mu\text{M/h}$		PGT	AICARFT	DHFR	TYMS	SHMT		Binding of 5 mTHF and SHMT	Unbinding of 5 mTHF and SHMT
						CH2F \rightarrow THF	THF \rightarrow CH2F		
CC	replete folate	5,268.8	6,749.8	263.4	263.4	15,646.6	4,367.7	14,964.5	1,4964.5
	low folate	4,711.3	6,473.4	117.8	117.8	18,535.4	7,268.2	8150.7	8150.7
absolute difference (% of replete folate)		10.6	4.1	55.3	55.3	18.5	66.4	45.5	45.5
TT	replete folate	5,475.7	6,843.0	322.1	322.1	66,523.2	39,415.5	6,559.7	6,559.7
	low folate	4,943.9	6,593.4	120.0	120.0	73,586.0	51,629.1	5,992.1	5,992.1
absolute difference (% of replete folate)		9.7	3.6	62.8	62.8	10.6	31.0	8.7	8.7

Table 9. Fluxes of the reactions catalyzed by the enzymes PGT, AICARFT, DHFR, TYMS and SHMT and the binding/unbinding of 5 mTHF and SHMT for replete (18 μM) and low (9 μM) levels of total folate and for the CC and TT case of the C677T MTHFR polymorphism.

also highly sensitive to folate status in the CC homozygotes, whereas GNMT flux in TT homozygotes was insensitive to folate status (Table 10). Similar but less pronounced effects were seen for flux through SHMT (Table 9).

To understand if *MTHFR* genotype affects the stability of the FOCM network at steady state, the deterministic simulation was coupled with stochastic simulation using a hybrid simulation strategy (see Methods and Supplementary Material). Model steady states were obtained under four different conditions that differed by *MTHFR* 677 genotype (CC and TT case) and intracellular folate levels (replete and low). Interestingly, the most stable steady state (the one with lowest total sum of reaction propensities $a_0(x)$, see Methods for details), was the CC case with folate replete concentrations, consistent with numerous epidemiological studies associating the *MTHFR* C677T genotype with folate-related pathologies (Table 11)⁴³. The enzyme that exhibited the greatest level of stochasticity in response to folate levels and/or the *MTHFR* C677T polymorphism was SHMT1 (Table S6).

The molecular basis of uracil misincorporation into DNA. Mouse models implicate SHMT1 and impaired *de novo* dTMP synthesis in NTD risk. Impaired *de novo* dTMP synthesis causes an increase in dUMP, which when converted to dUTP causes uracil misincorporation into DNA because DNA polymerases do not distinguish between dTTP and dUTP⁴⁴. The dTMP biosynthesis pathway enzymes (MTHFD1, SHMT, TYMS, and DHFR) are present in both the cytosol and recently have been found to function in the nucleus. In the nucleus, they comprise a multi-enzyme complex at sites of DNA synthesis that may be critical to limit rates of uracil misincorporation into DNA, but regulatory mechanisms remain unknown⁴⁵. These enzymes are modified by the Small Ubiquitin-like MOdifier (SUMO) protein at the G1/S boundary, which permits their nuclear translocation during S-phase of the cell cycle⁴⁶. One study showed that when nuclear translocation of this complex is impaired

$\mu\text{M/h}$		BHMT	MAT-I	MAT-III	GNMT	DNMT	SAHH	
							SAH \rightarrow HCY	HCY \rightarrow SAH
CC	replete folate	161.0	123.0	59.6	370.9	87.2	285.2	102.6
	low folate	170.3	128.8	55.3	664.0	51.8	294.0	110.0
absolute difference (% of replete folate)		5.8	4.7	7.3	79.0	40.5	3.1	7.2
TT	replete folate	175.5	129.9	52.7	830.4	34.7	297.1	114.5
	low folate	177.4	129.8	51.8	884.4	29.4	297.9	116.3
absolute difference (% of replete folate)		1.1	0.0	1.8	6.5	15.2	0.3	1.6

Table 10. Fluxes of the reactions catalyzed by the enzymes BHMT, MAT-I, MAT-III, GNMT, DNMT and SAHH for replete (18 μM) and low (9 μM) levels of total folate and for the CC and TT case of the C677T MTHFR polymorphism.

$a_0(x)$	replete folate	low folate	Difference (% of replete folate)
CC	$1.6054 \cdot 10^{15}$	$2.0970 \cdot 10^{15}$	30.62
TT	$8.5915 \cdot 10^{15}$	$1.0141 \cdot 10^{16}$	18.03
Difference (% of CC)	435.15	383.59	

Table 11. Total propensities obtained in four steady state conditions according to folate polymorphism (CC and TT case) and total concentration of available folate (replete, 18 μM ; low, 9 μM). To help comparisons, the differences between CC and TT (in % of CC) and between replete and low total folate (in % of replete folate) are indicated. A steady state that is less stable (or more noisy) than another one has higher total propensity.

in a mouse model over-expressing SHMT1, rates of uracil misincorporation into DNA increased several fold⁴⁵. In this model, SHMT1 protein levels were elevated several fold in the liver, yet its localization was restricted to the cytoplasm and nuclear SHMT1 levels were depleted compared to wild-type mice⁴⁵. Furthermore, nuclei isolated from SHMT1 overexpressing mice exhibited lower rates of *de novo* dTMP synthesis compared to nuclei isolated from wild-type mice⁴⁵. This suggests that *de novo* dTMP synthesis occurs when the enzymes are present in the multi-enzyme complex within the nucleus in mammals. However, no definitive experiment has been performed that identifies the relative contribution of nuclear and cytosolic dTMP synthesis to overall dTMP synthesis. Interestingly, *S. cerevisiae* do not import the dTMP synthesis pathway into the nucleus⁴⁷.

To determine if nuclear import of the *de novo* dTMP pathway was required to meet cellular demands for dTTP during DNA replication, rates of dTMP synthesis were modeled for mammalian cells using standard Michaelis-Menten kinetics (Table 12; Table 4). Based on the number of A-T base pairs in the human genome and an 8-hour S-phase in embryonic stem cells (S-phase in L1210 cells is also 6–10 h^{48, 49}), the rate of dTMP synthesis required for faithful cell replication is 7.8 $\mu\text{M}/\text{min}$ (calculations are in Supplementary Material, see also Table 12)^{50, 51}. In the current model, which does not account for SHMT1/TYMS/DHFR/MTHFD1 nuclear localization nor complex formation, cytosolic dTMP synthesis rates are 4.4 $\mu\text{M}/\text{min}$ (Table 4, DHFR and TYMS flux, 263.4 $\mu\text{M}/\text{h}$, assuming MTHFR 677 CC genotype). This computational deficit between dTMP requirements and dTMP synthesis rates suggests that dTMP synthesis as currently modeled in the cytosol where the enzymes are not present in a complex cannot meet cellular needs. Nuclear localization and complex formation of the *de novo* dTMP synthesis complex seem to be unique to mammalian cells. In *S. cerevisiae*, TYMS is not SUMOylated and localizes to the nuclear periphery⁴⁷. The measured rate of dTMP synthesis in *S. cerevisiae* is 1.8 $\mu\text{M}/\text{min}$ ⁵² (Table 12). The rate of dTMP synthesis required to replicate the *S. cerevisiae* genome over the course of an S-phase (less than one hour^{52, 53}) is 0.5 $\mu\text{M}/\text{min}$, indicating that yeast synthesize dTMP at a rate that is more than 3-fold greater than necessary for adequate dTMP synthesis (calculations are in Supplementary Material). Furthermore, in response to DNA damage, yeast increase dNTP concentrations 6–8 fold⁵⁴ and *E. coli* increase dNTP concentrations 1.8–3.7 fold⁵⁵, but dNTP concentrations do not increase after DNA damage in mammals^{56, 57}.

Discussion

Understanding the dynamics of FOCM and its responsiveness to both genetic and environmental perturbations is the key to understanding the etiology of folate-related pathologies. Computational models and related simulations permit an identification of the most sensitive reactions within the network that exhibit the greatest degree of stochastic behavior leading to variability in network outputs. Furthermore, computational models allow an understanding of how both genetics and environmental factors can enhance or repress stochastic behavior at defined locations within the network, accelerating the development of diagnostics to identify those at risk for folate-related pathologies as well as lead to the development of targeted nutritional interventions for disease prevention.

The *Shmt1* knockout mouse model (*Shmt1*^{+/-}, *Shmt1*^{-/-} embryos) exhibits impaired *de novo* dTMP synthesis in the absence of perturbations of homocysteine remethylation. It also recapitulates risk for NTDs in humans.

	Human	<i>S. cerevisiae</i>
Genome size	3.0×10^9 bp	1.2×10^7 bp
% AT	59%	61.5%
T bases needed for replication	1.77×10^9 molecules	7.5×10^6 molecules
Length of cell cycle	24 h	2.5 h
Length of S-phase	8 h	0.83 h
Cell volume	8×10^{-13} L (ES cell)	5×10^{-14} L
dTMP synthesis rate required to replicate genome	7.8 μ M/min	0.5 μ M/min
Measured dTMP synthesis rate (model outcomes)	4.4 μ M/min	1.8 μ M/min
Ratio of dTMP production relative to dTMP required for replication	0.6	3.6

Table 12. Cellular capacity for *de novo* dTMP synthesis in mammals and yeast at S-phase.

Specifically, the mouse model exhibits folate-responsive NTDs that occur with minimal perturbation in FOCM, and exhibit low and variable penetrance^{33,34}. In fact, most if not all, folate-related pathologies whose etiology involves interactions among genetic risk variants and nutrient exposures also exhibit low and/or variable clinical presentation. Understanding the stochastic behavior of the various reactions within FOCM that results in increased variability in FOCM network outputs is essential to understand which enzymes in the network contribute to folate-related pathologies.

Existing FOCM models rely on the limited quantity of kinetic data present in the literature, and the performance of the model will be dependent upon the kinetic parameters chosen to include in the model. Much of the available kinetic data for FOCM enzymes present in the literature was collected using the commercially available monoglutamate folate substrates, with few studies using the physiologically relevant polyglutamate forms of the cofactor. In the cell, newly transported monoglutamate folates are converted to folate polyglutamates, containing 3 to 7 polyglutamate moieties, though the action of folypolyglutamate synthetase⁵⁸. The polyglutamate chain (N = 3 glutamate and higher) increases the affinity of folate cofactors for many folate-dependent enzymes by one to two orders of magnitude^{39,40}. Models that include kinetic parameters derived from the use of folate monoglutamates can limit model reliability. Here we established a hierarchy of criteria to select a more homogeneous set of kinetic parameters (i.e. K_m and V_{max}) by referring, when possible, to L1210 cells because of the richness and quality of the data used to derive kinetic parameters. Furthermore, our preference was to select kinetic coefficients generated using folate polyglutamate cofactors and purified proteins from animal models closest to humans, as the variability in kinetic parameters among mammals is much less than the differences observed between folate monoglutamate and polyglutamate cofactor substrates.

The current model was validated by demonstrating that it recapitulates empirical observations regarding the impact of intracellular glycine on behavior of the FOCM network (Fig. 2). The validated model was then used to understand how the MTHFR C677T polymorphism, a known genetic risk factor for NTDs in humans, affects FOCM. This model shows that the lower levels of 5mTHF associated with the MTHFR 677 T variant are accompanied by elevated levels of 10fTHF, which has been observed in animal models and in humans^{27–29} (Table 1). The model also indicates that 10fTHF accumulates in TT homozygotes as a result of increased flux through both the synthetase activity of MTHFD1 (FTS activity, Table 3), but also due to increased flux through MTHFD1 activity in the direction converting CH2F to 10fTHF (Table 3). Therefore, the model accurately predicts perturbations in FOCM that have been observed in human clinical and epidemiological studies. A recent study suggested that the risk of the MTHFR C677T polymorphism for NTDs was due to its known effect on lowering intracellular folate concentrations, rather than its role in providing 5 mTHF for homocysteine remethylation³². This model demonstrates that the MTHFR C677T polymorphism elevates levels of 10fTHF, which is known to be a chemically unstable form of folate that is susceptible to oxidative degradation, providing a mechanism by which the MTHFR C677T polymorphism depletes intracellular folate levels. Importantly, the model reported here demonstrates that the *de novo* dTMP biosynthesis enzymes are the most sensitive to low intracellular folate concentrations, with both DHFR and TYMS activities being repressed by 63% (Table 9). This finding is consistent with the finding that mouse models with impaired *de novo* dTMP biosynthesis are susceptible to NTDs in folate deficiency³⁴. The hybrid stochastic simulation also reveals that both folate deficiency and the MTHFR C677T polymorphism create overall instability in the network (Table 11), consistent with a vast body of literature demonstrating an association of both folate deficiency and the MTHFR C677T polymorphism with various pathologies^{24,30}. Interestingly, SHMT1 exhibits the greatest increase in stochastic behavior as a result of the MTHFR C677T polymorphism (Table S6); the SHMT1 enzyme is the only FOCM enzyme that when disrupted results in folate-responsive NTDs³⁴.

The primary findings of this study are that the FOCM network is destabilized as a result of folate deficiency and the MTHFR677T polymorphism, and that SHMT is the most sensitive enzyme within the network to this network instability. This finding nicely connects the MTHFR genetic variant, a known risk factor for human NTDs, and SHMT1, the only folate enzyme whose disruption results in folate-responsive NTDs in mice. Furthermore, this model predicts that *de novo* dTMP synthesis rates in mammals are about half of what is required to meet DNA replication demands for dTMP (Table 12). Although mammals contain two pathways for dTMP synthesis,

the folate-dependent *de novo* dTMP synthesis pathway described here and a salvage pathway catalyzed by thymidine kinase 1, the salvage pathway activity is insufficient to meet cellular needs based on observations that folate deficiency results in elevated uracil accumulation in DNA. In mammalian cells, the *de novo* dTMP synthesis enzymes form a multi-enzyme complex that interacts with DNA replication enzymes⁴⁶. The discrepancy predicted by the model indicates that the model should be extended to include multi-enzyme complex formation and substrate channeling in the nucleus to model more accurately determinants of FOCM and dTMP synthesis. The inclusion of the dTMP multi-enzyme metabolic complex in the model is expected to limit substrate diffusion and increase the rate of dTMP synthesis.

Methods

Description of the Model and of the Simulation Techniques. The model was constructed as a closed system using the subset of reactions that describe the FOCM pathways and homocysteine remethylation in cytoplasm¹⁹ (Fig. 1). For the simulation, we employed a hybrid stochastic approach using deterministic simulation to compute the initial phase of the dynamics until a model steady state was reached, and then we assessed the stability of the achieved steady state by relying on exact stochastic simulation. We adopted a hybrid approach, rather than one entirely based on exact stochastic simulation⁵⁹, because of the intensive computational effort introduced by the stiffness of the system during the simulation.

The deterministic simulation was based on ODEs, where reactions were described in terms of Michaelis-Menten equations consistent with the original model of Reed¹⁹ and computed using the MATLAB integrator *ode15s*, whereas parameter estimates were derived from literature or calculated by nonlinear least squares optimization. The kinetic constants were obtained from folate polyglutamate cofactors and their interaction with enzymes purified from L1210 cells where possible and otherwise from other mammalian tissue (see Supplementary Material for a detailed discussion of the model and parameter estimates).

The set of ODEs was further translated into a stochastic reaction-based model and a hybrid simulation approach⁶⁰ was employed to quantify the level of stochasticity in the considered FOCM steady states. We refer to the Supplementary Material for a detailed description of how the ODE model has been translated to a reaction based stochastic one. According to the seminal work of Gillespie⁵⁹, exact stochastic simulation allows simulation of each reaction event asynchronously when such reaction event is most probable to occur. To allow this, the simulation algorithm computes a propensity function $a_j(x)$ at each simulation step for each modeled reaction R_j , where x is the current state of the system providing the abundances of all modeled species at the considered time. The propensity value of a reaction has a direct link with the probability of its execution, that is, reactions with higher propensity are more likely to be fired in the near future. To evaluate when the next reaction event will occur, also the total sum of propensities $a_0(x) = \sum_R a_j(x)$ is computed, because this quantity is linked to the number of reaction events occurring in the next time unit, that is, with increasing total propensity the number of reaction events per unit of time also increases. In Table 11, we provide the total sum of propensities for the considered steady states. On average, we would need to generate up to 10^{15} – 10^{16} reaction events per unit of time during the stochastic simulation, due to stiffness of the system. To circumvent the problem, we relied on the concept of total propensity to evaluate the stability of the steady state, by assuming that a steady state is more stable when it has a lower value of $a_0(x)$, that is, a lower averaged number of reaction events that can perturb the equilibrium of the steady state.

Model validation. To validate the FOCM model, *in silico* experiments were performed to determine if the model could recapitulate empirical data generated in MCF-7 cells by Herbig *et al.*¹⁰ focusing on the effect of glycine on FOCM (Fig. 2). Glycine is important because, as a second substrate, it has a direct influence on the reaction catalyzed by the enzymes GNMT as well as on the reversible reaction transforming CH₂F to THF catalyzed by the enzyme SHMT. The purpose is to understand how the steady state of FOCM is affected by altering intracellular glycine concentrations. This was achieved by running several model simulations starting from different glycine concentrations and comparing the corresponding steady states with empirical data from Herbig *et al.*¹⁰. This study examined the effect of exogenous glycine at concentrations from 0 to 10 mM on the relative distribution of folate one-carbon forms as well as S-adenosylmethionine (SAM) and S-adenosylhomocysteine (SAH) levels. In summary, the empirical data revealed that as glycine concentrations increase intracellular 10fTHF levels increase at the expense of 5mTHF levels that decrease. Furthermore, as glycine concentrations increase SAM levels are depleted and SAH levels rise. These changes were interpreted by the effects of glycine concentration driving the reversible SHMT reaction in the direction of serine synthesis¹⁰.

We simulated the effect of glycine on folate distribution, SAM, and SAH concentrations using the computational model for values of glycine ranging from 0 to 10 mM (Fig. 2). The trends obtained by the model simulations were in agreement with the literature (green arrows in Fig. 2), confirming the coherence between model outcomes and empirical data. We observed only one exception related to the total % of 5mTHF at 10 mM glycine (red arrow in Fig. 2). This discrepancy could be mainly due to two reasons: (1) 10 mM glycine is an extreme and non-physiological intracellular glycine concentration that could cause pharmacological effects, (2) the large magnitude of the experimental error in the Herbig *et al.* study at this glycine concentration.

References

1. Beaudin, A. E. & Stover, P. J. Insights into metabolic mechanisms underlying folate-responsive neural tube defects: a minireview. *Birth Defects Res A Clin Mol Teratol* **85**, 274–284 (2009).
2. Xu, X. *et al.* Folic acid therapy delays the progression of chronic kidney disease: the renal sub2 study of the China Stroke Primary Prevention Trial (CSPPT). *JAMA Internal Medicine* (2016).

3. Zhao, M., Li, X., Xing, C. & Zhou, B. Association of methylenetetrahydrofolate reductase C677T and A1298C polymorphisms with colorectal cancer risk: A meta-analysis. *Biomed Rep* **1**, 781–791 (2013).
4. Price, A. J. *et al.* Circulating Folate and Vitamin B12 and Risk of Prostate Cancer: A Collaborative Analysis of Individual Participant Data from Six Cohorts Including 6875 Cases and 8104 Controls. *Eur Urol* (2016).
5. Burr, N. E., Hull, M. A. & Subramanian, V. Folic Acid Supplementation May Reduce Colorectal Cancer Risk in Patients With Inflammatory Bowel Disease: A Systematic Review and Meta-Analysis. *J Clin Gastroenterol* (2016).
6. Visentin, M., Zhao, R. & Goldman, I. D. The antifolates. *Hematol Oncol Clin North Am* **26**, 629–648, ix (2012).
7. Ducker, G. S. & Rabinowitz, J. D. One-Carbon Metabolism in Health and Disease. *Cell metabolism* (2016).
8. Stover, P. J. Physiology of folate and vitamin B12 in health and disease. *Nutr Rev* **62**, S3–12, discussion S13 (2004).
9. Fox, J. T. & Stover, P. J. Folate-mediated one-carbon metabolism. *Vitam Horm* **79**, 1–44 (2008).
10. Herbig, K. *et al.* Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses. *J Biol Chem* **277**, 38381–38389 (2002).
11. Scotti, M., Stella, L., Shearer, E. J. & Stover, P. J. Modeling cellular compartmentation in one-carbon metabolism. *Wiley interdisciplinary reviews. Systems biology and medicine* **5**, 343–365 (2013).
12. Seither, R. L., Trent, D. F., Mikulecky, D. C., Rape, T. J. & Goldman, I. D. Folate-pool interconversions and inhibition of biosynthetic processes after exposure of L1210 leukemia cells to antifolates. Experimental and network thermodynamic analyses of the role of dihydrofolate polyglutamylates in antifolate action in cells. *J Biol Chem* **264**, 17016–17023 (1989).
13. Jackson, R. C. & Harrap, K. R. Studies with a mathematical model of folate metabolism. *Arch Biochem Biophys* **158**, 827–841 (1973).
14. Morrison, P. F. & Allegra, C. J. Folate cycle kinetics in human breast cancer cells. *J Biol Chem* **264**, 10552–10566 (1989).
15. Strong, W. B., Tendler, S. J., Seither, R. L., Goldman, I. D. & Schirch, V. Purification and properties of serine hydroxymethyltransferase and C1-tetrahydrofolate synthase from L1210 cells. *J Biol Chem* **265**, 12149–12155 (1990).
16. Martinov, M. V., Vitvitsky, V. M., Mosharov, E. V., Banerjee, R. & Ataullakhanov, F. I. A substrate switch: a new mode of regulation in the methionine metabolic pathway. *J Theor Biol* **204**, 521–532 (2000).
17. Reed, M. C., Nijhout, H. F., Sparks, R. & Ulrich, C. M. A mathematical model of the methionine cycle. *J Theor Biol* **226**, 33–43 (2004).
18. Prudova, A., Martinov, M. V., Vitvitsky, V. M., Ataullakhanov, F. I. & Banerjee, R. Analysis of pathological defects in methionine metabolism using a simple mathematical model. *Biochim Biophys Acta* **1741**, 331–338 (2005).
19. Reed, M. C. *et al.* A mathematical model of glutathione metabolism. *Theor Biol Med Model* **5**, 8 (2008).
20. Nijhout, H. F. *et al.* A mathematical model gives insights into the effects of vitamin B-6 deficiency on 1-carbon and glutathione metabolism. *J Nutr* **139**, 784–791 (2009).
21. Reed, M. C. *et al.* A mathematical model gives insights into nutritional and genetic aspects of folate-mediated one-carbon metabolism. *J Nutr* **136**, 2653–2661 (2006).
22. Goss, P. J. & Peccoud, J. Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets. *Proc Natl Acad Sci USA* **95**, 6750–6755 (1998).
23. Gillespie, J. H. Multilocus Behavior in Random Environments. II. Linkage Disequilibrium in an Additive Model. *Genetics* **87**, 569–579 (1977).
24. Schwahn, B. & Rozen, R. Polymorphisms in the methylenetetrahydrofolate reductase gene: clinical consequences. *American journal of pharmacogenomics: genomics-related research in drug development and clinical practice* **1**, 189–201 (2001).
25. Chiang, E. P., Wang, Y. C. & Tang, F. Y. Folate restriction and methylenetetrahydrofolate reductase 677T polymorphism decreases adoMet synthesis via folate-dependent remethylation in human-transformed lymphoblasts. *Leukemia* **21**, 651–658 (2007).
26. Quinlivan, E. P. *et al.* Methylenetetrahydrofolate reductase 677C->T polymorphism and folate status affect one-carbon incorporation into human DNA deoxynucleosides. *The Journal of nutrition* **135**, 389–396 (2005).
27. Bagley, P. J. & Selhub, J. A common mutation in the methylenetetrahydrofolate reductase gene is associated with an accumulation of formylated tetrahydrofolates in red blood cells. *Proc Natl Acad Sci USA* **95**, 13217–13220 (1998).
28. Davis, S. R. *et al.* Homocysteine synthesis is elevated but total remethylation is unchanged by the methylenetetrahydrofolate reductase 677C->T polymorphism and by dietary folate restriction in young women. *J Nutr* **135**, 1045–1050 (2005).
29. Ghandour, H., Chen, Z., Selhub, J. & Rozen, R. Mice deficient in methylenetetrahydrofolate reductase exhibit tissue-specific distribution of folates. *J Nutr* **134**, 2975–2978 (2004).
30. Stover, P. J., MacFarlane, A. J. & Field, M. S. Bringing clarity to the role of MTHFR variants in neural tube defect prevention. *Am J Clin Nutr* **101**, 1111–1112 (2015).
31. Solis, C. *et al.* Folate intake at RDA levels is inadequate for Mexican American men with the methylenetetrahydrofolate reductase 677TT genotype. *J Nutr* **138**, 67–72 (2008).
32. Tsang, B. L. *et al.* Assessing the association between the methylenetetrahydrofolate reductase (MTHFR) 677C > T polymorphism and blood folate concentrations: a systematic review and meta-analysis of trials and observational studies. *Am J Clin Nutr* **101**, 1286–1294 (2015).
33. Martiniova, L., Field, M. S., Finkelstein, J. L., Perry, C. A. & Stover, P. J. Maternal dietary uridine causes, and deoxyuridine prevents, neural tube closure defects in a mouse model of folate-responsive neural tube defects. *Am J Clin Nutr* **101**, 860–869 (2015).
34. Beaudin, A. E. *et al.* Shmt1 and de novo thymidylate biosynthesis underlie folate-responsive neural tube defects in mice. *Am J Clin Nutr* **93**, 789–798 (2011).
35. Beaudin, A. E. *et al.* Dietary folate, but not choline, modifies neural tube defect risk in Shmt1 knockout mice. *Am J Clin Nutr* **95**, 109–114 (2012).
36. Woeller, C. F., Anderson, D. D., Szebenyi, D. M. & Stover, P. J. Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway. *J Biol Chem* **282**, 17623–17631 (2007).
37. Field, M. S. *et al.* Nuclear enrichment of folate cofactors and methylenetetrahydrofolate dehydrogenase 1 (MTHFD1) protect de novo thymidylate biosynthesis during folate deficiency. *J Biol Chem* **289**, 29642–29650 (2014).
38. Field, M. S., Kamynina, E., Watkins, D., Rosenblatt, D. S. & Stover, P. J. Human mutations in methylenetetrahydrofolate dehydrogenase 1 impair nuclear de novo thymidylate biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 400–405 (2015).
39. Hilton, J. G., Cooper, B. A. & Rosenblatt, D. S. Folate polyglutamate synthesis and turnover in cultured human fibroblasts. *J Biol Chem* **254**, 8398–8403 (1979).
40. Schirch, V. & Strong, W. B. Interaction of folylpolyglutamates with enzymes in one-carbon metabolism. *Arch Biochem Biophys* **269**, 371–380 (1989).
41. Liew, S. C. & Gupta, E. D. Methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism: epidemiology, metabolism and the associated diseases. *European journal of medical genetics* **58**, 1–10 (2015).
42. Anguera, M. C. *et al.* Regulation of folate-mediated one-carbon metabolism by 10-formyltetrahydrofolate dehydrogenase. *J Biol Chem* **281**, 18335–18342 (2006).
43. Ueland, P. M., Nygard, O., Vollset, S. E. & Refsum, H. The Hordaland Homocysteine Studies. *Lipids* **36**(Suppl), S33–39 (2001).
44. Blount, B. C. *et al.* Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proc Natl Acad Sci USA* **94**, 3290–3295 (1997).
45. Macfarlane, A. J. *et al.* Nuclear localization of the De Novo thymidylate biosynthesis pathway is required to prevent uracil accumulation in DNA. *J Biol Chem* (2011).

46. Anderson, D. D., Woeller, C. F., Chiang, E. P., Shane, B. & Stover, P. J. Serine hydroxymethyltransferase anchors de novo thymidylate synthesis pathway to nuclear lamina for DNA synthesis. *J Biol Chem* **287**, 7051–7062 (2012).
47. Poon, P. P. & Storms, R. K. Thymidylate synthase is localized to the nuclear periphery in the yeast *Saccharomyces cerevisiae*. *The Journal of biological chemistry* **269**, 8341–8347 (1994).
48. Young, R. C. & DeVita, V. T. The effect of chemotherapy on the growth characteristics and cellular kinetics of leukemia L1210. *Cancer Res* **30**, 1789–1794 (1970).
49. Neil, G. L. & Homan, E. R. The effect of dose interval on the survival of L1210 leukemic mice treated with DNA synthesis inhibitors. *Cancer Res* **33**, 895–901 (1973).
50. Becker, K. A. *et al.* Self-renewal of human embryonic stem cells is supported by a shortened G1 cell cycle phase. *Journal of cellular physiology* **209**, 883–893 (2006).
51. Sagi, I. *et al.* Derivation and differentiation of haploid human embryonic stem cells. *Nature* **532**, 107–111 (2016).
52. Greenwood, M. T., Calmels, E. M. & Storms, R. K. Growth-rate-dependent regulation of the expression and inactivation of thymidylate synthase in *Saccharomyces cerevisiae*. *Journal of bacteriology* **168**, 1336–1342 (1986).
53. Brewer, B. J., Chlebowicz-Sledziewska, E. & Fangman, W. L. Cell cycle phases in the unequal mother/daughter cell cycles of *Saccharomyces cerevisiae*. *Mol Cell Biol* **4**, 2529–2531 (1984).
54. Chabes, A. *et al.* Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. *Cell* **112**, 391–401 (2003).
55. Gon, S., Napolitano, R., Rocha, W., Coulon, S. & Fuchs, R. P. Increase in dNTP pool size during the DNA damage response plays a key role in spontaneous and induced-mutagenesis in *Escherichia coli*. *Proc Natl Acad Sci USA* **108**, 19311–19316 (2011).
56. Hakansson, P., Hofer, A. & Thelander, L. Regulation of mammalian ribonucleotide reduction and dNTP pools after DNA damage and in resting cells. *The Journal of biological chemistry* **281**, 7834–7841 (2006).
57. Niida, H., Shimada, M., Murakami, H. & Nakanishi, M. Mechanisms of dNTP supply that play an essential role in maintaining genome integrity in eukaryotic cells. *Cancer science* **101**, 2505–2509 (2010).
58. Stover, P. J. & Field, M. S. Trafficking of intracellular folates. *Adv Nutr* **2**, 325–331 (2011).
59. Gillespie, D. T. Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem* **58**, 35–55 (2007).
60. Marchetti, L., Priami, C. & Thanh, V. H. Efficient hybrid stochastic simulation for spatially homogeneous biochemical reaction networks. *J. Comput. Phys.* **317**, 301–317 (2016).
61. Gostner, R., Baldacci, B., Morine, M. J. & Priami, C. Graphical Modeling Tools for Systems Biology. *ACM Comput. Surv.* **47**, 1–21 (2015).

Acknowledgements

We would like to acknowledge Stephen M Stover for technical assistance. We are grateful to Lorenzo Stella, Tommaso Schiavinotto, Chiara Damiani and Alessandro Romanel for their helpful suggestions and discussions. This work was supported by Public Health Service grant DK58144.

Author Contributions

K.M., L.M., M.S., C.P., M.S.F., P.J.S. contributed to the experimental design. K.M., L.M., M.S. developed the computation model, while M.S.F., P.J.S., and M.S. provided all biological insight and extracted kinetic variables from the literature. M.S.F., P.J.S., and C.P. provided overall guidance of the project. All authors contributed to manuscript preparation and approved the final version.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-00854-w](https://doi.org/10.1038/s41598-017-00854-w)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017